J. Clin. Chem. Clin. Biochem.
Vol. 17, 1979, pp. 57–64

Problems of Identification in the Clinical Laboratory, State of the Art and other Considerations[1])

By *H. Keller*

Institut für Klinische Chemie und Hämatologie des Kantons St. Gallen

(Received June 21/September 14, 1978)

Dedicated to Prof. Dr. Theodor Bücher on the occasion of his 65th birthday

Summary: In recent years, there have been no new commercial developments in identification methods for the clinical chemical laboratory. We are still faced, however, with the problem of ensuring accurate identification. By using a model, the importance of reliable identification can be shown mathematically. The complete identification of a sample is based on a five point identification statement. As a rule, appropriate measures are taken to abridge this identification process, depending on the local situation. The terminology of identification processes is defined and the presently available techniques are discussed.

Probleme der Identifikation im klinischen Laboratorium, Stand der Technik und weitere Überlegungen

Zusammenfassung: In den letzten Jahren wurden keine industriellen Neuentwicklungen für Identifikationsprozesse im klinisch-chemischen Laboratorium angeboten. Die Schwierigkeiten der Identifikations-Sicherung bestehen aber unverändert weiter. An einem Modell kann die Bedeutung der sicheren Identifikation rechnerisch belegt werden. Zur vollständigen Probenidentifikation dient ein aus fünf Punkten zusammengesetztes Identifikations-Statement. Es wird durch entsprechende Hilfsmittel und Maßnahmen in der Regel für innerbetriebliche Zwecke verkürzt. Die für Identifikationsprozesse gebräuchlichen Begriffe werden definiert und die heute verfügbaren Techniken diskutiert.

Preliminary Remarks

Publications dealing with the importance and the difficulties of identification processes in the clinical chemistry laboratory are relatively few (1–6). Within the framework of scientific conferences, this theme was last discussed at the "Biochemische Analytik" meeting in 1972 (7, 8), and in the same year in Munich at the 8th World Pathology Congress (9).

A review of the present state of the technology reveals that there have been no specific new developments in identification processes in the clinical chemical laboratory in the last six years. All the currently used identification techniques were already known in 1972.

The same applies to instrumentation. Product promotion has claimed most attention, whereas fundamental new developments in instrumentation have only very occasionally appeared, and there has been no willingness to exploit them.

The importance of unequivocal identification in clinical work is unquestionable. It must be possible to guarantee that false negative, or false positive results are not produced by mistakes in information transfer. What is the situation in the general medical laboratories, where most of the results are in the so-called normal range? One might favour the view that it does not matter if the analytical profiles of two patients, both showing no abnormalities, are interchanged. Since false identification also includes re-exchange, the mistakes could even cancel themselves.

Since, despite the technical feasibility, longitudinal observations are hardly ever carried out, there would appear to be a certain justification for this view.

Model Systems

Increasing interest has been shown in recent years in the use of the probability theory in medical diagnosis (10); it may also be applied to the problem of sample identification. This is illustrated by the first example (fig. 1):

[1]) Revised from a paper given by the author at the international meeting "Biochemische Analytik 78" held in Munich on 18th to 21st April 1978.

0340-076X/79/0017-0057$2.00

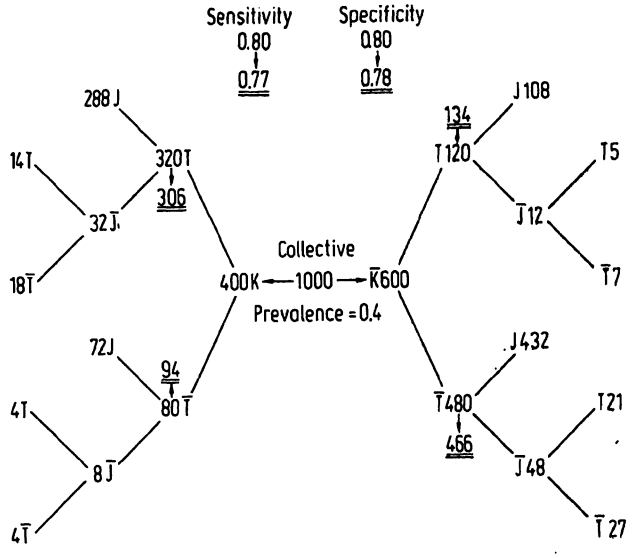

Fig. 1. Decision network; symbols:
K = patients
$\bar{K}$ = non-patients
T = positive test results
$\bar{T}$ = negative test result
J = correct identification
$\bar{J}$ = false identification
The doubly underlined numbers show the final result. For further details, see the text.

We will take a collective of patients and healthy individuals; these are to be differentiated with the aid of a single test, which simply yields a negative or positive result. We will also assume that the prevalence of the disease is 0.4, and the sensitivity and specificity are 0.8. The certainty of identification is 0.9, so that in 100 samples 90 are identified correctly, but 10 are identified wrongly.

The decision network shows the influence of insufficient identification on the sensitivity of the test. Out of 400 patients, 320 show a correct positive test result, and 80 show a false negative result. These proportions are altered yet again by the mistakes of identification. Thus, of the 320 genuine positives, 288 are correct, and 32 are wrongly identified, of the 80 false negatives, 72 are correct and 8 are wrongly identified. For the 40 wrongly identified samples, there are two further possibilities: Faulty identification can lead to a positive or negative test result. The 32 wrongly identified (but true positive) samples now divide into 14 positive and 18 false negatives. Of the 8 false negatives, 8 are wrongly identified, and they then divide into 4 false positives and 4 positives which are now correct. The result is therefore 306 positives and 94 negatives. The binary test result gives no decision as to whether the correctly positive result was obtained by a correct or false identification. The analogous study of the specificity, which is likewise assumed to be 0.8, shows that the 600 healthy individuals give 480 correctly negative and 120 falsely positive results. By wrong identification, a further 5 false positives are now added, while 7 are changed back again to correct negatives. There occur 48 identification errors in the 480 true negatives, whereby 27 true negatives remain, while 21 are changed to false positives. The calculation of the probability that a re-exchange will occur is shown in the example of the negative test results (fig. 2). In summary, it can be ascertained that mistaken identification leads to a decrease in the number of true positives, while the false negatives are increased. This means that the sensitivity is effectively decreased from 0.8 to 0.77, and the specificity from 0.8 to 0.78.

If the prevalence is decreased to 0.1 (fig. 3), the sensitivity decreases from 0.8 to 0.75, whereas the specificity only decreases from 0.8 to 0.79.

This leads to the following conclusion: the lower the prevalence of the disease in the collective, the more serious is the effect of wrong identification on the sensitivity of the test. In other words the chance of identifying a clinically abnormal person in a collective of many healthy individuals is further decreased.

$$\Sigma T = 320 + 120 = 440$$

$$\Sigma \bar{T} = 80 + 480 = 560$$

$$1000$$

$$P(\bar{T}|\bar{J}) = \frac{\bar{T}}{\bar{T} + T} = \frac{560}{1000} = 0.56$$

Fig. 2. Calculation of the probability of a false identification with a negative test result. For further details, see the text.

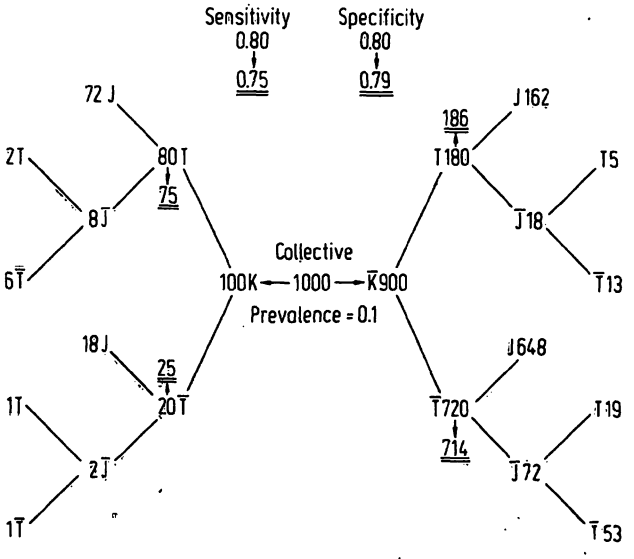

Fig. 3. Decision network as in figure 1. . .

Definitions

Sample identification

The term *sample identification* (in the clinical chemical laboratory) embraces all those processes and/or measures that serve to establish an identification statement for a given quantity of test material from a patient. Such a statement consists of the following five factors:

1. *Source of the sample*

   As a rule, this is the name of the patient from whom the sample was taken. This can be supplemented with information that is useful for interpretation purposes: age, sex, diagnosis/condition, therapeutic measures, pretreatment before taking the sample, etc.

2. *Source of the request*

   This is usually a doctor or clinic. Additional details may include, for example, a request for communication of the results by telephone, or a requirement to send the results to more than one person or institution, etc.

3. *Test material*

   It must be clear whether the material is serum, plasma with additives, blood with or without additives, or other test material with or without added foreign substances.

4. *Instructions*

The instructions must consist of:

a) a statement of the required test or tests.
b) data required for the evaluation and calculation of the results, e.g. 24-hour urine volume, or time of blood sampling following glucose loading.

5. *Timing*

The following must be recorded:

a) the time or period of the day and the date that the sample was taken from the patient.
b) The date and time that the material was received by the laboratory.

This complete identification statement is necessary, if the patient and test sample are separated in distance and time from the laboratory and the analysis. The statement must be made permanent, e.g. written by hand, or recorded on a data carrier. An exception can be made, if the investigation is carried out directly on the patient. For example, if the patient is continually monitored by an ECG machine (and the apparatus also records the time), the identification statement can be omitted, identification processes are not required, and it is not even necessary to know the name of the patient.

Direct sample identification

In *direct sample identification*, each specimen and each process vessel can be described at any time in relation to the identification statement. At every stage that involves a measurement, the appropriate value is processed together with the corresponding identification in a data set. Two basically different procedures are possible:

1. *Direct sample identification without aids to identity*

   The sample container and identification statement are kept together from the time of sampling to the conclusion of the analysis in the laboratory. The results of all the measurements performed are entered on the statement, so that they cannot be lost or confused.

   Such conditions are, however, the exception. In clinical laboratories this process is occasionally used in the determination of blood gases: The sample container is the syringe, containing the freshly sampled arterial blood, and this is brought immediately to the laboratory. Blood gases are then determined accordingly in an ear-marked apparatus. The results are written directly onto the accompanying statement sheet; and they are often then communicated by telephone.

2. *Direct sample identification with aids to identity*

   In routine operation, the complete identity statement is mostly too extensive and contains more information than is required for the analytical processes. For this reason, the statement is usually abbreviated, with the aid of a code, to an *identification symbol.* This may consist of a patient number, day number and/or other relevant numbers, e.g. the nature of the test or tests required.

Specimen vessels, process vessels and analytical records are linked to one another by the identification symbols. Retrieval of the complete identification statement necessitates a decoding process, which may be based on, for example, a "day list". Most simply, a manually prepared list is used, while in better appointed laboratories, these lists are stored in an electronic data processor (EDP).

Direct sample identification does not include the time, place or nature of the identification process. Neither does it guarantee that there is a positive link between sample and measurement signal and/or result. In the latter situation, i.e. if a EDP is installed and the identification is checked at every stage from the moment of distribution to the production of the result, and if the simultaneous mechanical transfer of the identification symbol is compulsory at every transfer step, then the procedure is justifiably called *permanent identification*, a term introduced by *Eggstein* (11).

When the specimen and process vessels can be identified, by virtue of carrying day numbers, patient numbers or similar data, which are not, however, read visually or mechanically at every working stage and processed with the result, then the procedure can only be called *partial direct identification.*

Indirect sample identification

There are two variants of *indirect identification*: in *positional* identification, the sample is identified from its position in a chain of samples, a magazine, or sample plate etc., which, in turn, is related to a list. Identification may take place, for example, by counting the members of the chain, whereby regular empty spaces act as markers. If only the order of samples, together with a work list is used for identification, then the procedure is called *serial identification.* This latter method carries the greatest danger of sample interchange in the laboratory.

A *code* is a syntactical working instruction. It shows, for example, which position corresponds to a certain number. The intellectual process of translating a several digit number into a code is the *coding*; this is normally performed mechanically. In any event, it is a process that depends upon software. The corresponding hardware counterpart performs the *marking*, e.g. the punching of holes in a card. The holes punched in a particular pattern are known as *marks.* Analogous processes are the pencil-marking of labelling sheets, or the printing of other information carriers, that serve for identification.

Although the definitions are clear, there is often a failure, in english, to distinguish clearly between these different procedures by use of the correct terminology.

The extent of influencing factors

Identification processes depend upon three factors:

1. The quantity of the specimen and the number of different tests (range of the laboratory program).
2. Instrumentation and organisation of the operational schedule.
3. Distance between laboratory and patient.

Each factor has a different effect on the nature and scope of the identification process.

A general laboratory that processes 1000 samples per day on five days of the week, and only uses a fixed program on a multichannel analyser, can manage with very simple identification procedures.

A small clinical laboratory may have a low degree of mechanisation, but a large test program and discriminating analysis. It must therefore use a sample distribution system, in which the identification processes function just as smoothly as each subsequent analytical step. Here, the question of patient admission may represent a serious problem. Patient admission is frequently an urgent matter, yet test results without a corresponding set of controls can lead to uncertainties.

A study of methods for the interpretation of instructions, sample distribution and identification in the large clinical laboratory, was recently published by *Killian* (12), in which he employed the methods of operational research. It is assumed that the samples arrive in the laboratory in exponentially distributed intervals. In addition, the laboratories always have a redundant range of apparatus, i.e. the 20 most important tests can always be performed on at least two different instruments. In his model, *Killian* uses an adjoint matrix and develops waiting queue-algorithms, which are based on *Poisson*-distributed rates of arrival. The redundant instruments are used in order to fulfill the optimal requirements for economy and speed. Such complex distribution processes, however, presuppose not only a functional EDP, but an equally highly developed identification system. This requires a much greater outlay for identification purposes than in the first example.

The experience of recent years has shown the identification process within the laboratory depends essentially upon the instrumentation. The *central sample distribution* with central documentation which is still typically used in most clinical laboratories, is economically inferior to the *consecutive sample distribution* and documentation. Admittedly, central sample distribution is not usually practised consistently in the clinical laboratory. Some analysers are suited to a consecutive distribution technique, e.g. the SMAC (Technicon), the ACA (Du Pont), the GSAII (Greiner) and the PRISMA (Clinicon), but their test programs alone are not sufficient for the larger hospital laboratories. These large laboratories have to provide special system configurations to meet special requirements (for example the determination of blood glucose), while at the same time meeting the clinical requirements of a much larger test program. Consecutive sample distribution, whose identification process is determined by the available analyser, cannot therefore be practised throughout the laboratory. Also, many difficult analyses cannot be completed in one day, so that identification must be extended over a fairly long period. This considerably increases the difficulty of ensuring the correctness of the identification process.

A minimum of identification is necessary in bedside analysis, or "monitoring". In the last six years, however, the proportion of clinical-chemical bedside analyses has remained low and unchanged.

Identification Techniques:

Various techniques are used for direct identification. The most common of these are given in figure 4.

Eppendorf

The Eppendorf labelling system (13) is based on reflection marks, that are imprinted on the process vessels. In the completely planned laboratory, the required analyses can be related to the available apparatus at the outset, and all the necessary process vessels for one

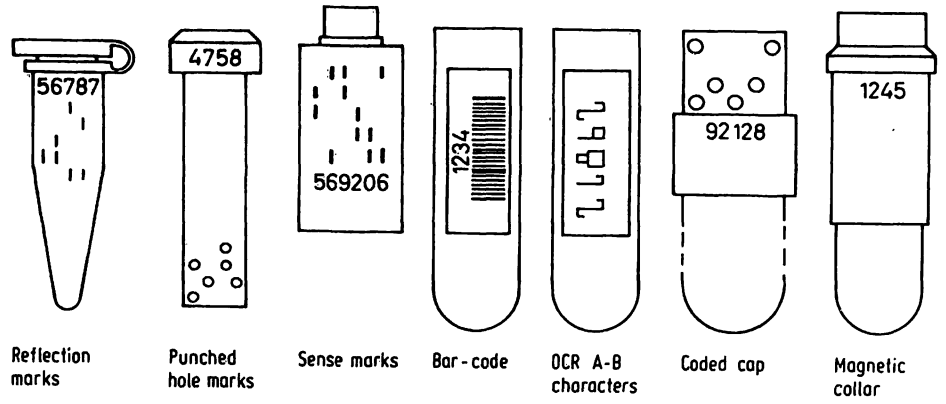

Fig. 4. Methods of direct sample identification.

sample can be labelled at the mechanical sample distribution site. The marks, which are read automatically by the machine, represent patient or day numbers, and they correspond to a five channel binary code. They are supplemented by visual characters, so that an operator check is possible at any time. This particular identification system is widely used by laboratories in central Europe.

SILAB

The SILAB-System represents a similar philosophy, but it employs a different labelling technique (14, 15, 16). The excellent reading-reliability of this system is unquestionable. Unfortunately, the further development and sale of this identification system have been discontinued.

Technicon

Label-reading by machine was realised by Technicon in the IDEE-system (17, 18, 19), and it has been retained in a modified form. This method is especially widely used in Great Britain, but it is also popular in other European laboratories.

IBM

Forerunner of this system was the IBM 1080 data system with perforated punched cards which was introduced by *Rappoport* (20, 21, 22) in the USA, and which is still in use in many American and a few European laboratories. There have been many interesting modifications of the perforated punched card, e.g. by *Borner* (23), *Jentzsch* (24) and others. The book „Diagnostik-Informations-System" still offers many useful ideas and suggestions (25); the situation up to 1975 has been reported by *Haeckel* (26).

Bar code

Nearly 10 years ago, *Rubin* (27) drew attention to the Bar-code. Almost simultaneously, *Dudek & Roka* (28) also reported satisfactory results with this labelling system. Following a period of stagnation, renewed interest is now being shown in this technique. In 1977, *Laue* (29) reported detailed studies with the Bar-code: the identification labels are not restricted to any one kind of vessel; in addition to specimen tubes, they can also be used for haematocrit capillaries, urine vessels, blood sugar tubes etc. Labels and analysis requests carry the Bar-code label, but they can also be read visually. It appears that the success of this system is due to the fact that the reading process is especially simple; it can be performed rapidly and reliably, either manually with the aid of a reading pistol. or mechanically with a laser beam. It should be pointed out that the reading unit, which is required in order to achieve high reading reliability, is relatively cheap. With the aid of the EDP, Bar-code labels are produced for the primary serum vessel and for all the other necessary process vessels. These labels carry the identification number of the patient, together with the requested type of investigation. Experience has shown that the reading reliability is not endangered even by badly smudged labels. The system has, however, the following disadvantages:

1. The labels are so large that they cannot be stuck onto microlitre vessels.
2. The durability of the marking instruments is limited, and
3. The label printers have a relatively high rate of breakdown and they are expensive.

American blood banks were recently recommended by the American Blood Commission (30) to label all blood storage vessels with Bar-code labels, and to introduce the CODABAR reading system. It is to be expected that this type of identification, at least in the USA, will soon be used very widely. Meanwhile, Technicon have fitted their blood group analyser with a Bar-code-laser-scanner (fig. 5). Reading by the laser beam requires no contact with the label; the labels can also be read at an angle, and they can even be stuck in a slanting position on the sample. The Bar-code system has also been fitted into two large analysers, which are so far in the prototype stage, namely the Hycel M and the PRISMA (Clinicon).

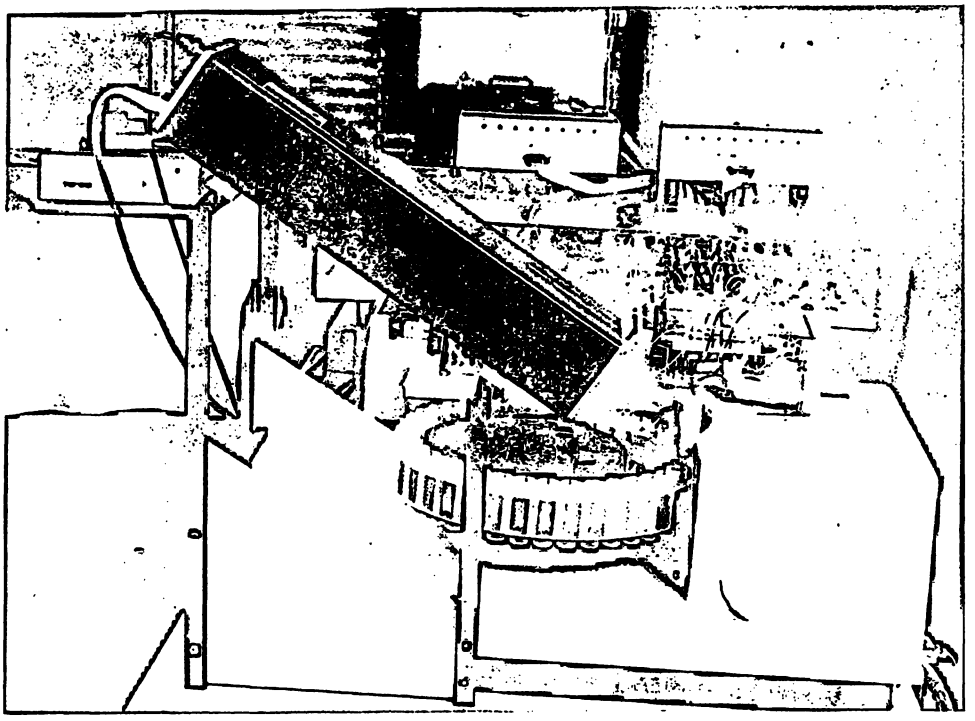

Fig. 5. Prototype for reading Bar-code with the aid of a Laser beam.

OCR

The most recent and perhaps most promising development is optical character recognition (OCR) (31). Optional character sets for OCR at present in common use for single scanners are shown in figure 6. Provisionally, it is possible to read every digit, although a more economically priced reading apparatus will limit the reading ability to the first 10 letters. Operation of the reading pistol, shown here for the Eppendorf ELIS system, is extremely simple. Its error-free and relatively quick use comes easily to the hands of even non-specialised personel (fig. 7).

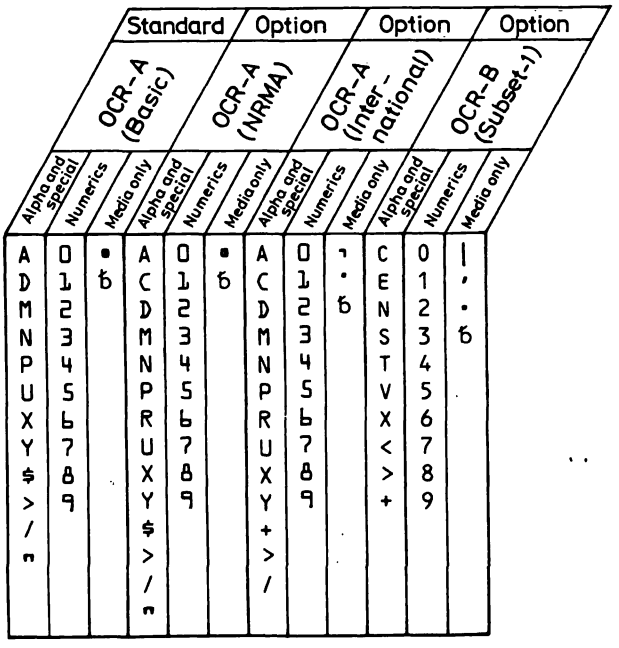

Fig. 6. Optional character sets for OCR.

OCR has the following advantages:

1. OCR is a direct code, and the information can be read both automatically and visually.
2. No special printing apparatus is needed; the labels and other information carriers can be prepared for OCR with any golf-ball typewriter, or EDP-controlled rapid printer.
3. In the forseeable future, OCR systems will be able to read automatically the name of the patient and other alpha-numerical information.
4. OCR labelling and reading are becoming increasingly popular in the commercial field. Further development should therefore be accelerated, and the prices of the reading apparatus should come down.
5. The central institute for contributory health insurance ("kassenärztliche Versorgung") in the German Federal Republic has developed a method for assessing medical services, which is also based on the OCR system (32). This could also contribute to decreasing the costs and to shortening the time for development. Disadvantages of the simple OCR reader at the moment are the limited number of symbols and the relatively slow manual operation of the reading pistol. In big laboratories, where several hundred samples must be read in a short time, a scanner for mark-sense cards is at present essential.

Others

The following two possibilities for sample identification have seen no further development:

The coded cap, planned for the SILAB system (33), acted as a seal for the specimen vessel, and its edge represented a ring-type, five channel system of strip holes. It carried data for patient number, process number

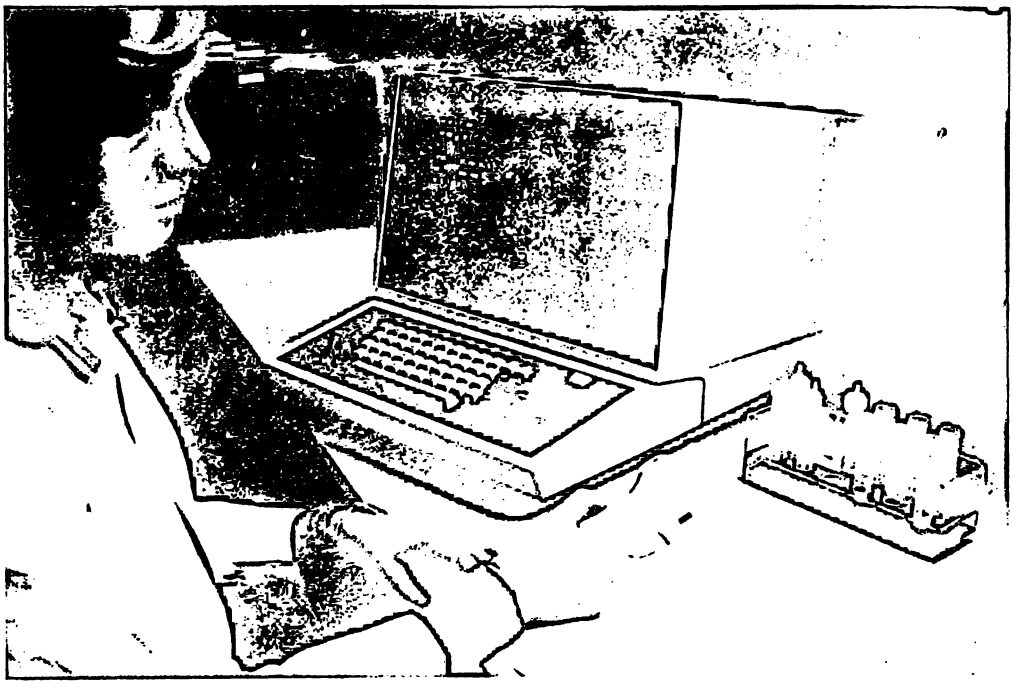

Fig. 7. Operation of the OCR reading pistol (Eppendorf ELIS system).

and/or the requested analytical method, and it served as the information basis for the sample distributer. The failure to adopt this elegant and imaginative concept was certainly to a large extent due to economic considerations.

Another original concept was the magnetic collar described by *Grabner* et al (34) in 1974. This consisted of an identification ring containing six steel wires, which could be magnetised to give an 18-bit information system. Static reading was by the *Hall* effect. Unfortunately this system also found no acceptance. Equally unsuccessful were the magnetic caps fitted to the Mark X 10 produced by the firm Hycel.

Final Observations

From the foregoing discussion, it is clear that in all methods presently employed, it is not the material under investigation that is identified with respect to the source of the material but the sample container and/or chain of containers. The source of material and the corresponding analytical results can only be related by a reconstruction process. Thus the weakest link in the identification chain is revealed: The unequivocal identification of the patient and the reliable initial transfer of the patient identification to the specimen (i.e. specimen-vessel). None of the suggestions for the improvement of the reliability of this decisive step, e.g. an identification carrier permanently associated with each patient, has found wider application (35). Whenever the taking or collection of samples is not performed by laboratory personel, all further efforts to ensure correct identification are fraught, from the outset, with a non-calculable uncertainty factor.

In summary, it must be said that in the last six years, there has been no significant progress in the field of identification of patient, specimen or process vessel; although most large laboratories now have some form of electronic data processing. The expectation, that the introduction of EDP's would greatly influence the identification process, has not been fulfilled.

The Bar-code or the OCR system appear to have the best chances for the future. With both systems, however, there are still certain technical difficulties and as yet there is no sign that these will be overcome.

At the present time, it is therefore not possible to make general, firm recommendations for identification systems in clinical chemical laboratories. In the absence of any alternative for ensuring the reliability of every identification process, the stability of the operation and organisation of the laboratory will always be crucial. Together with manifold technical and scientific aids, moderation, accuracy and self criticism remain, now as ever, the most important parameters in our laboratories; indeed they are more important in the patient identification process.

References:

1. Delbrück, A. & Hock, D. (1969), Krankenhausarzt *42*, 308–312.
2. Blatt, E. (1972), Datenverarbeitung im klinisch-chemischen Laboratorium. G.I.T. Verlag Darmstadt.
3. Keller, H. (1972), Diagnostik *5*, 273–276.
4. Keller, H. (1972), Diagnostik *5*, 277–280.
5. Keller, H. (1972), Diagnostik *7*, 320–324.
6. Blatt, E. (1978), Biomed. Technik *23*, Ergänzungsband Mai/Juni 1978. Beitrag No. 100.
7. Büttner, H., Eggstein, M., Keller, H. & Laue, D. (1972), Mitt. Dtsch. Ges. Klin. Chem. 2/72, 29–43.
8. Keller, H. (1972), Z. Klin. Chem. Klin. Biochem. *10*, 482–488.
9. Keller, H. & Richterich, R. (1972), Proc. VIII World Congress of Anatomic and Clinical Pathology Munich 12–16. Sept. 1972. Excerpta Medica Amsterdam. Internat. Congr. Series No 285, 183–189.
10. Büttner, J. (1977), J. Clin. Chem. Clin. Biochem. *15*, 1–12. Dort auch weitere Literatur.
11. Eggstein, M. (1978), Das Präsenzlabor. Springer-Verlag, Berlin (im Druck).
12. Killian, M. (1977), G.I.T. Fachz. Lab. *21*, 1218–1226.
13. Hille, W. & Knaus, M. (1973), G.I.T. Fachz. Lab. *17*, 1255.
14. Knedel, M., Schipper, Ph., Haas, W., Killian, K. & Krech, H. (1970), Vortrag Analytika München 1970, zitiert nach Blatt, E. l.c. (2).
15. Gräser, W., Mieth, J., Porth, A., Knodel, W. & Eggstein, M. (1969), G.I.T. Fachz. Lab. Sonderheft Lab. Medizin Mai 1969, 426–434.
16. Knedel, M. (1971), Siemens Druckschrift Nr MC 50/1016.
17. Junge, B. & Hoffmeister, H. (1974), Technicon Bibliography 8088.
18. Sher, P. P. & Gambino, S. R. (1970), Adv. Autom. Analysis Vol I, 63–65, Thurman Ass. Miami, Florida
19. Helmreich, W. (1969), Sample identification for automated analysis system. Technicon Internt. Congress June 4–6, 1969, Chicago, Ill.
20. Rappoport, A. E., Gennaro, W. D. & Constandse, W. J. (1968), Modern Hospital *110*, 100–103.
21. Rappoport, A. E., Gennaro, W. D. & Constandse, W. J. (1967), Med. Hospital *108*, 107–111.
22. Rappoport, A. E., Tom, Y. D., Gennaro, W. D. & Borquist, R. E. (1977), Damming the flood of laboratory data. Vortrag: 2nd Internat. Congr., Birmingham Sept 12–14, 1977.
23. Borner, K. & Klein, E. (1969), Z. Klin. Chem. Klin. Biochem. 7, 185–188.
24. Jentzsch, D. (1969), Probenidentifizierung bei klinisch-chemischen Analysen. Berichte zur klinischen Chemie. Schriftenreihe herausgegeben vom Bodenseewerk Perkin Elmer & Co., Ueberlingen/Bodensee, Heft 1 (1969).
25. Bock, H. E. & Eggstein, M. (eds.) (1970), Diagnostik Informationssystem, Springer Verlag, Berlin.
26. Haeckel, R. (1976), Rationalisierung des medizinischen Laboratoriums G.I.T. Verlag Darmstadt.
27. Rubin: (1969), VII Int. Congr. Clin. Chem. Geneva 1969, Diskussionsbeitrag.
28. Dudek, J., Roka, L. & Michel, H. (1974), Ärztl. Lab. *20*, 430–436.
29. Laue, D. (1977), The BAR-CODE: A universal system of identification in the medical laboratory. Birmingham 2nd Internat. Conference, Sept. 12–16, 1977.
30. American Blood Commission (1976), CCBBA special issue August 1976. New Blood pack label being tested. Am. Blood Commission, Arlington, Va. 22209.
31. Michel, H. (1977), Biotechn. Umschau *1*, 23.
32. Geiss, D. & Schwartz, D. (1977), Arzt und Wirtschaft *21*, 31–32.
33. Beckmann, V., Haas, W. & Killian, K. (1972), Diagnostik *6*, 284–286.
34. Marksteiner, A., Bancsich, J. & Grabner, H. (1974), Aerztl. Lab. *20*, 426–430.
35. Selander, H. (1970), Patient identification. In: Information Processing of Medical Records (Anderson, J. & Forsythe, J. M., eds.). North Holland Publ. Co. Amsterdam.
36. Müllertz, S. (1970), Scand. J. Clin. Lab. Invest. *26*, 407–413.

Prof. Dr. Dr. H. Keller
Inst. f. Klin. Chemie und Hämatologie
Frohbergstr. 3
CH-9000 St. Gallen